1) Why do we use Derivative / Gradient to fit a target function?

   Because gradient can make correct direction and spend less time to convergence.

2) In the words 'Gradient Descent', what's the Gradient and what's the Descent?

   Gradient is partial derivative. Descent is negative direction.

3) What's the advantages of the 3rd gradient descent method compared to the previous methods?

   The 3rd method can convergence more quickly and then save a lot of time.

4) Using the simple words to describe: What's the machine learning?

   Machine learning is a method of data analysis that automates analytical modeling building by using data.

5) Why do we need dynamic programming? What's the difference of dynamic programming and previous talked search problem?

   Dynamic programming can save last time's calculation result. Hence, it can save a lot of time and computer memory. Even more, computer cannot handle one question that has too many iterations and does not use dynamic programming.

   The difference between them is dynamic programming can save last time's calculation result.

6) Why do we still need dynamic programming? Why not we train a machine learning to fit a function which could get the right answer based on inputs?

   It can handle one question that has too many iterations.

   It will occur overfitting.

7) Can you catch up at least 3 problems which could solved by Dynamic Programming?
   1) Coin change, returning the min number of coins
   2) Coin change, returning the maximum change methods
   3) Shortest distance, returning the shortest distance from original node to target node

8) Can you catch up at least 3 problems which could solved by Edit Distance?
   1) Edit distance between 2 strings is one
   2) Delete operation for two strings (only deletion operation is used)
   3) Minimum ASCII Delete sum for two strings

9) Please summarize the three main features of Dynamic Programming, and make a concise explain for each feature.
   1) Optimal substructure: An optimal solution can be constructed from optimal solutions of its subproblems.
   2) Future states is based on previous states and don't affect previous states.
   4) Overlapping problems: A recursive algorithm for the problem solves the same subproblem over and over rather than always generating new subproblems

10) What's the disadvantages of Dynamic Programming? (You may need search by yourself in Internet)
   1) No general formation of dynamic programming is available, every problem has to be solving in its own way.
   2) Dividing problem in sub problem and storing inter mediate results consumes memory.